UNIVERSIDADE FEDERAL DE SANTA CATARINA

LIMITES SUPERIOR E INFERIOR DE CAPACIDADE GENERALIZADA

Marli Carlesso

Agosto - 1983

LIMITES SUPERIOR E INFERIOR DE CAPACIDADE GENERALIZADA

por

Marli Carlesso

Esta dissertação foi julgada adequada para a obtenção do título de

"MESTRE EM CIÊNCIAS"

Especialidade em Matemática, e aprovada em sua forma final pelo Curso de Pós-Graduação em Matemática da

UNIVERSIDADE FEDERAL DE SANTA CATARINA

Prof. William Glenn Whitley Ph.D.
Coordenador

Banca Examinadora:

Prof. Inder Jeet Taneja, Ph.D.
Orientador

Prof. Augusto C. Gadelha Vieira, Ph.D.

Prof. Gur Dial, Ph.D.

RESUMO

Arimoto (1972) e Blahut (1972) apresentaram o procedimento iterativo para calcular a capacidade de um canal discreto sem memória. Arimoto (1972) também apresentou os limites superior e inferior da capacidade do canal discreto sem memória. Em 1975 e 1976, Arimoto novamente apresentou um procedimento iterativo para cálculo da capacidade generalizada de ordem ρ, C_ρ, tendo por base a entropia de ordem $\alpha (\alpha = \frac{1}{1+\rho})$, relacionando-a com uma função importante do teorema de codificação (Gallager (1965)), chamada "função exponencial de codificação aleatória".

O objetivo principal deste trabalho é apresentar pela primeira vez na literatura matemática, os limites superior e inferior de capacidade generalizada de ordem ρ, C_ρ. Estes são apresentados no Capítulo 2. No Capítulo 1, apresentamos alguns conceitos e resultados básicos para o desenvolvimento deste trabalho, e no final daremos alguns valores numéricos dos resultados obtidos, para três canais diferentes.

ABSTRACT

Arimoto (1972) and Blahut (1972) presented an iterative procedure to calculate the capacity of a discrete memoryless channel. Arimoto (1972) also presented upper and lower bounds on the capacity of a discrete memoryless channel. In 1975 and 1976, Arimoto again presented an iterative procedure for generalized capacity of order ρ, C_ρ, using entropy of order $\alpha (\alpha = \frac{1}{1+\rho})$ and related it to the important function of coding theorem (Gallarger (1965)), called "randon coding exponent function.

The main aim of this work is to present for the first time in the mathematic literature, the upper and lower bounds on the generalized capacity of order ρ, C_ρ. This we have done in Chapter II. In Chapter I, we presented some basic concepts and results in order to develop this work. At the end, we give some numerical values of all the results obtained for three different channels.

ÍNDICE

CAPÍTULO 1

INTRODUÇÃO

O problema fundamental da Teoria de Informação é quantificar o conteúdo informacional de uma mensagem. Esta mensagem é propagada através do canal de comunicação, sendo a capacidade do canal conhecida como o máximo de informação mútua entre as entradas e saídas que pode ser transmitida através dele, onde o máximo é tomado sobre todas as distribuições de probabilidade de entrada.

Em 1948, Shannon apresentou o seguinte teorema: "Dado um canal ruidoso, com capacidade C, é possível transmitir mensagens sobre este e ainda decodificá-las com uma probabilidade de erro arbitrariamente pequena, desde que a taxa, ou velocidade de transmissão, seja menor que C e que as mensagens possam ser codificadas em blocos de comprimento N.
Entretanto se a taxa de transmissão exceder C, a probabilidade de erro tende a 1 quando N tende ao infinito"

A seguir, apresentamos alguns conceitos e resultados básicos para o desenvolvimento do nosso trabalho.

1.1 - Informação Mútua

Definição 1.1.1

Seja $X = \{1,2,...,n\}$ uma variável aleatória discreta, com distribuição de probabilidade $p = (p_1,p_2,...,p_n)$ onde p pertence ao conjunto P^n de vetores probabilidade $P^n = \{p \in R^n / p_i \geq 0; \sum_{i=1}^{n} p_i = 1\}$ e $Y = \{1,2,\ldots,m\}$ uma variável aleatória discreta, com distribuição de probabilidade $q = (q_1,q_2,\ldots,q_m)$; $q \in P^m$.

A informação mútua entre X e Y é definida por:

$$I(X/Y) = H(X) - H(X/Y) \tag{1.1}$$

onde H(X) é a entropia de Shannon e H(X/Y) é a entropia condicional de X dado Y, dadas respectivamente por:

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i$$

e

$$H(X/Y) = -\sum_{i=1}^{n} \sum_{j=1}^{m} p(i,j) \log p(i/j)$$

Podemos escrever

$$I(X/Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_i\, p(j/i) \log \frac{p(j/i)}{q_j} \tag{1.2}$$

onde $q_j = \sum_{i=1}^{n} p_i\, p(j/i)$, $j = 1,2,\ldots,m$

Esta informação mútua satisfaz as seguintes propriedades: (referência Ash (1965))

(i) $I(X/Y) \geq 0$ (não negativa) (1.3)

(ii) $I(X/Y) = I(Y/X)$ (simétrica) (1.4)

(iii) $I(X/Y)$ é uma função côncava sobre P^n (1.5)

Obs.: Consideramos os logarítmos na base 2.

1.2 - Canal Discreto Sem Memória

Definição 1.2.1

Um canal discreto sem memória com alfabeto de entrada $X = \{1,2,...,n\}$ e alfabeto de saída $Y = \{1,2,...,m\}$, com distribuição de entrada $(p_1,p_2,...,p_n)$ e distribuição de saída $(q_1,q_2,...,q_m)$ é caracterizado pela matriz n x m

$$\{p(j/i)\} = \begin{bmatrix} p(1/1) & p(2/1) & \dots & p(m/1) \\ p(1/2) & p(2/2) & \dots & p(m/2) \\ \text{---} & \text{---} & \text{---} & \text{---} \\ p(1/n) & p(2/n) & \dots & p(m/n) \end{bmatrix} \tag{1.6}$$

onde $p(j/i)$, $i = 1,2,...,n$; $j = 1,2,...,m$ representa a probabilidade condicional de receber $Y = j$ enquanto $X = i$ é transmitida; com $\sum_{j=1}^{m} p(j/i) = 1$; $i = 1,2,...,n$

A matriz $\{p(j/i)\}$ é chamada matriz-canal

Definição 1.2.2 - (Capacidade do Canal)

A capacidade C de um canal discreto sem memória é definida como o valor máximo de $I(X/Y)$, onde a maximização é tomada sobre o conjunto de todas as distribuições de probabilidade de entrada $p \in P^n$, isto é,

$$C = \max_{p \in P^n} I(X/Y) \tag{1.7}$$

O teorema a seguir é demonstrado, e.g. em Gallager (1968), p. 87-88.

Teorema 1.2.3

Um conjunto de condições necessária e suficiente para um vetor probabilidade de entrada $p = (p_1, p_2, \ldots, p_n)$ atingir a capacidade C de um canal discreto sem memória é que:

$$I(X = i/Y) \begin{cases} = C & \text{se } p_i > 0 \quad \forall_i,\ i = 1,2,\ldots,n \\ \leq C & \text{se } p_i = 0 \quad \forall_i,\ i = 1,2,\ldots,n \end{cases} \tag{1.8}$$

onde $I(X = i/Y)$ é a informação transmitida pelo canal, para entrada $X = i$ a partir da saída Y, dada por:

$$I(X = i/Y) = \sum_{j=1}^{m} p(j/i) \log \left(\frac{p(j/i)}{\sum_{k=1}^{n} p_k\, p(j/k)} \right) \tag{1.9}$$

para todo $i = 1,2,\ldots,n$

Agora é apresentada a definição de probabilidade de erro, cujos limites superior e inferior são apresentados a seguir.

Sejam:

X^N o conjunto de todas as seqüências de entrada, de comprimento N, que podem ser transmitidas, isto é,

$X^N = \{x = (x_1 x_2 \ldots x_N)$ tal que $x_t \in X$ para $t = 1,2,\ldots N$; X finito$\}$

Y^N o conjunto de todas as seqüências de saída de comprimento N que pode ser recebida, isto é, $Y^N = \{y = (y_1 y_2 \ldots y_N)$ tal que $y_t \in Y$ para $t = 1,2,\ldots,N$; Y finito$\}$

Definição 1.2.4 - Probabilidade de Erro

A probabilidade de erro de decodificação para um certo código B, com M palavras código de comprimento N, denotado por P_e é definida por:

$$P_e = \sum_{x_m \in X^N} \frac{1}{M} P_{em} \tag{1.10}$$

onde P_{em} é a probabilidade de erro de decodificação, quando a seqüência x_m é transmitida.

Quando $x = (x_1 x_2 \ldots x_N) \in X^N$ é transmitida, a probabilidade de receber a seqüência $y = (y_1 y_2 \ldots y_N) \in Y^N$ é determinada por:

$$p(y/x) = \prod_{t=1}^{N} p(y_t / x_t) \tag{1.11}$$

Em 1965 Gallager apresentou uma demonstração simples do teorema de codificação de Shannon e obteve novo limite superior de probabilidade de erro de decodificação P_e.

Este limite para taxa $R = \frac{\log M}{N}$, onde M é o número de palavras código de comprimento N, de um certo código B é dado por:

$$P_e \leq \exp\left[-N\left\{-\rho R + \max_{p} E_o(\rho,p)\right\}\right] \tag{1.12}$$

$$\text{para} \quad 0 \leq \rho \leq 1$$

Arimoto (1973) apresentou um limite inferior da seguinte forma:

$$P_e \geq 1 - \exp[-N\{-\rho R + \min_{p} E_o(\rho,p)\}] \qquad (1.13)$$

$$\text{para} \quad -1 \leq \rho < 0$$

onde

$$E_o(\rho,p) = -\log[\sum_{j=1}^{m} \{\sum_{i=1}^{n} p_i \, P(j/i)^{\frac{1}{1+\rho}}\}^{1+\rho}] \qquad (1.14)$$

sendo $P = \{p(j/i)\}$ uma matriz probabilidade de transição que caracteriza um canal discreto sem memória e p uma distribuição de probabilidade do alfabeto de entrada $X = \{1,2,\ldots,n\}$. A maximização e minimização de $E_o(\rho,p)$ em (1.12) e (1.13) é tomada sobre o conjunto de todas as distribuições de probabilidade de entrada. Ainda em 1973, Omura estendeu o limite inferior da probabilidade de erro P_e para $R > C$ com $-\infty < \rho \leq 0$.

Foi observado por Gallager (1968) que o problema da maximização da função exponencial de codificação aleatória $E_o(\rho,p)$, é quase idêntico ao da obtenção da capacidade. Mas, por outro lado, Arimoto (1975), relaciona a função exponencial de codificação aleatória $E_o(\rho,p)$ com a capacidade de ordem ρ, definindo a informação mútua em termos da entropia de ordem α (Renyi (1961)), onde $\alpha = \frac{1}{1+\rho}$ e, em (1976) ele apresentou um método iterativo para computá-la.

A entropia de ordem α, à qual me referi é apresentada à seguir.

1.3 - Entropia de Ordem α

Renyi (1961) propôs uma generalização da entropia H(X), proposta por Shannon, denominada entropia de ordem α, que para uma distribuição de probabilidade $(p_1,p_2,\ldots,p_n) \varepsilon P^n$ de uma variá vel aleatória discreta $X = \{1,2,\ldots,n\}$ é definida por:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \{ \sum_{i=1}^{n} p_i^\alpha \}$$

ou

$$H_\alpha(X) = \frac{\alpha}{1-\alpha} \log \{ \sum_{i=1}^{n} p_i^\alpha \}^{\frac{1}{\alpha}} \tag{1.15}$$

$$\alpha > 0 \quad ; \quad \alpha \neq 1$$

sendo que $\lim\limits_{\alpha \to 1} H_\alpha(X) = H(X)$ (1.16)

A entropia condicional de ordem α de X dado Y é então definida co mo: (Arimoto (1975))

$$H_\alpha(X/Y) = \frac{\alpha}{1-\alpha} \log \left[\sum_{j=1}^{m} \{ \sum_{i=1}^{n} p_i^\alpha \, p^\alpha(j/i) \}^{\frac{1}{\alpha}} \right] \tag{1.17}$$

$$\text{onde} \quad \alpha > 0 \quad , \quad \alpha \neq 1$$

Com esta definição

$$\lim_{\alpha \to 1} H_\alpha(X/Y) = H(X/Y) \tag{1.18}$$

e

$$H_\alpha(X/Y) \leq H_\alpha(X) \tag{1.19}$$

com igualdade se e somente se X e Y são independentes.

<u>Definição 1.3.1</u> - Informação de ordem α transmitida pe lo canal.

A informação de ordem α sobre X transmitida pelo canal a partir de Y é dada por:

$$I_\alpha(X/Y) = H_\alpha(X) - H_\alpha(X/Y)$$

$$= \frac{-\alpha}{1-\alpha} \log \left[\frac{\sum_{j=1}^{m} \{ \sum_{i=1}^{n} p_i^\alpha \, p^\alpha(j/i) \}^{\frac{1}{\alpha}}}{\{ \sum_{i=1}^{n} p_i^\alpha \}^{\frac{1}{\alpha}}} \right] \tag{1.20}$$

onde $\alpha > 0 \quad ; \quad \alpha \neq 1$

sendo $\lim_{\alpha \to 1} I_\alpha(X/Y) = I(X/Y)$ (1.21)

1.3.2 - Capacidade de Ordem α de Um Canal Discreto Sem Memória

A capacidade de ordem α de um canal discreto sem memória é definida como sendo o máximo de $I_\alpha(X/Y)$ tomado sobre todas as distribuições de probabilidade de entrada $p = (p_1, p_2, \ldots, p_n) \in P^n$, isto é,

$$C_\alpha = \max_{p \in P^n} I_\alpha(X/Y) \tag{1.22}$$

Devemos observar que $I_\alpha(X/Y)$ pode ser escrito como:

$$I_\alpha(X/Y) = \frac{-\alpha}{1-\alpha} \log \left[\sum_{j=1}^{m} \{ \sum_{i=1}^{n} p_i' \, p(j/i)^\alpha \}^{\frac{1}{\alpha}} \right] \tag{1.23}$$

$$\alpha > 0 \quad ; \quad \alpha \neq 1$$

onde $p_i' = \dfrac{p_i^\alpha}{\sum_{k=1}^{n} p_k^\alpha} \quad ; \quad i = 1, 2, \ldots, n$

Logo, $I_\alpha(X/Y) = \frac{\alpha}{1-\alpha}\tilde{E}_o(\alpha, p')$

onde

$$\tilde{E}_o(\alpha, p') = -\log\left[\sum_{j=1}^{m}\left\{\sum_{i=1}^{n} p'_i\, p(j/i)^{\alpha}\right\}^{\frac{1}{\alpha}}\right] \qquad (1.24)$$

com $p' = (p'_1, p'_2, \ldots, p'_n) \in P^n$ um vetor probabilidade de entrada.

A função $\tilde{E}_o(\alpha, p')$ é equivalente à função $E_o(\rho, p)$, como dada em (1.14) com $\alpha = \frac{1}{1+\rho}$, $\frac{1}{2} \leq \alpha < \infty$ e portanto

$$I_\alpha(X/Y) = \frac{1}{\rho} E_o(\rho, p) \quad , \quad \rho = \frac{1-\alpha}{\alpha} \quad , \quad \alpha > 0 \qquad (1.25)$$

Vale então os seguintes resultados (Arimoto (1975)).

Lema 1.3.3

1) $C_\alpha = \max_{p \in P^n} \frac{\alpha}{1-\alpha}\tilde{E}_o(\alpha, p) \qquad (1.26)$

onde $\tilde{E}_o(\alpha, p)$ é equivalente a $E_o(\rho, p)$ com a convenção $\rho = \frac{1-\alpha}{\alpha}$.

2) C_α é uma função contínua de α em $[\frac{1}{2}, \infty)$.

3) $\lim_{\alpha\to 1} C_\alpha = C \qquad (1.27)$

4) C_α é uma função não decrescente de α crescente.

O teorema a seguir, apresenta uma relação entre a probabilidade de erro e a capacidade de ordem α.

Teorema 1.3.4 (Arimoto (1975; 1976))

1) Existe pelo menos um código de comprimen o N, tal

que a probabilidade de erro de decodificação é limitada superiormente por:

$$P_e \leq \exp\ [N\ \{\frac{\alpha-1}{\alpha}\ (C_\alpha - R)\}] \tag{1.28}$$

$$\text{para } \frac{1}{2} \leq \alpha < 1$$

2) Para qualquer código de comprimento N a probabilidade de erro de decodificação é limitada inferiormente por:

$$P_e \geq 1 - \exp\ [N\ \{\ \frac{\alpha-1}{\alpha}\ (C_\alpha\ - R)\}] \tag{1.29}$$

$$1 < \alpha < \infty$$

O resultado acima portanto, reduz os limites superior e inferior da probabilidade de erro P_e de (1.12) e (1.13) respectivamente para (1.28) e (1.29).

Em (1976), Arimoto apresentou um algorítmo iterativo para calcular o máximo de $E_o(\rho, p)$, visto em (1.14), definindo a capacidade de ordem ρ, C_ρ como o máximo duplo da função $G_\rho(p, \Phi)$, conforme dado a seguir (ver (1.44)).

1.4 - Maximização da função exponencial $E_o(\rho, p)$ para Canal Discreto Sem Memória.

A função $G_\rho(p, \phi)$ é definida por:

$$G_\rho(p, \Phi) = - \frac{1}{\rho} \log\ [\sum_{j=1}^{m} \sum_{i=1}^{n} p_i\ p(j/i)\ \{\frac{\Phi(i/j)}{p_i}\}^{-\rho}] \tag{1.30}$$

onde $p = (p_1, p_2, \ldots, p_n) \in P^n$ e Φ pertence ao conjunto de matrizes transição probabilidade $n \times m$, tal que

$$\Phi^{nxm} = \{\Phi(./.) \text{ tal que } \Phi(i/j) \geq 0 \; ; \; \sum_{i=1}^{n} \Phi(i/j) = 1\} \tag{1.31}$$

A matriz Φ^{nxm} é geralmente chamada "dummy backward channel".

Para $\rho = 0$ em (1.30), é verificado facilmente que:

$$G_o(p, \Phi) = \lim_{\rho \to o} G_\rho(p, \Phi) = \sum_{j=1}^{m} \sum_{i=1}^{n} p_i \, p(j/i) \log \frac{\{\Phi(i/j)\}}{p_i} \tag{1.32}$$

O procedimento iterativo para calcular a capacidade generalizada de ordem ρ, à qual já nos referimos, tem por base o seguinte teorema.

Teorema 1.4.1 (Arimoto (1976))

a) Para algum $\rho \in (-1, \infty)$ fixado e algum $p \in P^n$ fixado, $G_\rho(p, \Phi)$ é maximizado por:

$$\Phi(i/j) = \frac{p_i \; p(j/i)^{\frac{1}{1+\rho}}}{\sum_{k=1}^{n} p_k \; p(j/k)^{\frac{1}{1+\rho}}} \tag{1.33}$$

e então

$$G_\rho(p) = \max_{\Phi} G_\rho(p, \Phi) = -\frac{1}{\rho} \log \left[\sum_{j=1}^{m} \{ \sum_{i=1}^{n} p_i \; p(j/i)^{\frac{1}{1+\rho}} \}^{1+\rho} \right] \tag{1.34}$$

b) Para algum $\rho \in (-1, \infty)$ fixado, e algum $\Phi \in \Phi^{nxm}$ fixado

$G_\rho(p, \Phi)$ é maximizado por:

$$p_i = \frac{\{\sum_{j=1}^{m} p(j/i)\ \Phi(i/j)^{-\rho}\}^{-\frac{1}{\rho}}}{\sum_{k=1}^{n} \{\sum_{j=1}^{m} p(j/k)\ \Phi(k/j)^{-\rho}\}^{-\frac{1}{\rho}}} \tag{1.35}$$

Então

$$G_\rho(\Phi) = \max_p G_\rho(p,\Phi) = \log \left[\sum_{i=1}^{n} \{ \sum_{j=1}^{m} p(j/i)\ \Phi(i/j)^{-\rho} \}^{-\frac{1}{\rho}} \right] \tag{1.36}$$

<u>Observação</u>:

É conveniente observar no entanto, que para $\rho = 0$

$$\Phi(i/j) = \frac{p_i\ p(j/i)}{\sum_{k=1}^{n} p_k\ p(j/k)} \tag{1.37}$$

e (1.35) torna-se

$$p_i = \frac{\exp\left[\sum_{j=1}^{m} p(j/i) \log \Phi(i/j)\right]}{\sum_{k=1}^{n} \exp\left[\sum_{j=1}^{m} p(j/i) \log \Phi(k/j)\right]} \tag{1.38}$$

Como conseqüência direta de (1.33) e (1.34) observamos que

$$E_o(\rho, p) = \rho \max_\Phi G_\rho(p, \Phi) \tag{1.39}$$

Também por (1.32) e (1.33)

$$I(X|Y) = \lim_{\rho\to o} \max_{\phi} G_\rho(p,\phi) \tag{1.40}$$

Vale então os seguintes resultados

Corolário 1.4.2

$$\max_p E_o(\rho,p) = \rho[\max_p \max_\phi G_\rho(p,\phi)] \tag{1.41}$$

$$\text{para} \quad 0 < \rho < \infty$$

$$\min_p E_o(\rho,p) = \rho[\max_p \max_\Phi G_\rho(p,\Phi)] \tag{1.42}$$

$$\text{para} \quad -1 < \rho < 0$$

Sendo a maximização e minimização de $E_o(\rho,p)$ tomada sobre o conjunto de todas as distribuições de probabilidade de entrada.

$$C = \lim_{\rho\to o} [\max_p \max_\phi G_\rho(p,\phi)] \ , \tag{1.43}$$

onde C é a capacidade do canal discreto sem memória.

Em vista destas relações, a capacidade generalizada de ordem ρ, denotada por C_ρ, pode ser escrita em máximo duplo como:

$$C_\rho = \max_p \max_\Phi G_\rho(p,\phi). \tag{1.44}$$

ou seja, por (1.34),

$$C_\rho = \max_p G_\rho(p) = \max_p \{-\frac{1}{\rho}\log [\sum_{j=1}^{m} \{\sum_{i=1}^{n} p_i\, p(j/i)^{\frac{1}{1+\rho}}\}^{1+\rho}]\}$$

Como vemos (1.45)

$$C = \lim_{\rho\to 0} C_\rho \tag{1.46}$$

e ainda

$$\text{(i)} \quad \max_p E_o(\rho,\ p) = \rho C_\rho \tag{1.47}$$

para $0 < \rho < \infty$

$$\text{(ii)} \quad \min_p E_o(\rho,\ p) = \rho C_\rho \tag{1.48}$$

para $-1 < \rho < 0$

O lema seguinte dá condições necessárias e suficientes para um canal discreto sem memória atingir a capacidade generalizada.

<u>Lema 1.4.2</u> (Jelinek (1968), Guerra (1980))

Um conjunto de condições necessárias e suficientes para $p^o \varepsilon P^n$ maximizar $G_\rho(p) = \max_\Phi G_\rho(p,\Phi)$ é que

$$\begin{cases} -\frac{1}{\rho}\log [\sum_{j=1}^{m} p(j/i)^{\frac{1}{1+\rho}} \{\sum_{i=1}^{n} p_i^o\, p(j/i)^{\frac{1}{1+\rho}}\}^{\rho}] = C_\rho & (1.49) \\ \qquad \text{se } p_i^o > 0 & \\ -\frac{1}{\rho}\log [\sum_{j=1}^{m} p(j/i)^{\frac{1}{1+\rho}} \{\sum_{i=1}^{n} p_i^o\, p(j/i)^{\frac{1}{1+\rho}}\}^{\rho}] \leq C_\rho & (1.50) \\ \qquad \text{se } p_i^o = 0 & \end{cases}$$

A seguir é apresentado um procedimento iterativo para calcular a capacidade de ordem ρ (Arimoto (1976))

1.5 - Procedimento Iterativo e Convergência

O procedimento iterativo é composto dos seguintes passos:

(i) Escolha um vetor probabilidade arbitrário $p^1 \varepsilon P^n$

É conveniente na prática, tomar uma distribuição uniforme de probabilidade $p_i^1 = \frac{1}{n}$ para todo $i = 1,2,\ldots,n$.

As seguintes etapas são iterativas para $t = 1,2,\ldots$

(ii) Para algum $\rho \varepsilon (-1, \infty)$ maximize $G_\rho(p^t, \Phi^t)$ com respeito a $\Phi \varepsilon \Phi^{nxm}$, fixado p^t. O valor de Φ que maximiza G_ρ é dado por

$$\Phi^t(i/j) = \frac{p_i^t \, p(j/i)^{\frac{1}{1+\rho}}}{\sum\limits_{k=1}^{n} p_k^t \, p(j/k)^{\frac{1}{1+\rho}}} \tag{1.51}$$

(iii) Para algum $\rho \varepsilon (-1, \infty)$ maximize $G_\rho(p, \Phi)$ em relação a $p^t \varepsilon P^n$, para todo $i = 1,2,\ldots,n$, fixando Φ^t

Este vetor probabilidade, denotado por p^{t+1}, é dado por:

$$p_i^{t+1} = \frac{\left[\sum\limits_{i=1}^{n} p(j/i) \, \{\Phi(i/j)\}^{-\rho}\right]^{-\frac{1}{\rho}}}{\sum\limits_{k=1}^{n} \left[\sum\limits_{j=1}^{m} p(j/k) \, \{\Phi(k/j)\}^{-\rho}\right]^{-\frac{1}{\rho}}} \tag{1.52}$$

Em vista do teorema (1.4.1) observa-se que este processo de cálculo consiste de maximizações sucessivas com respeito a $p \varepsilon P^n$ e $\Phi \varepsilon \Phi^{nxm}$.

Com base nestas maximizações sucessivas, Arimoto (1976), obteve os seguintes resultados:

Lema 1.5.1

$$G_\rho(p^1, \Phi^1) \le G_\rho(p^2, \Phi^1) \le G_\rho(p^2, \Phi^2) \le \dots \le G_\rho(p^t, \Phi^t)$$

$$\le G_\rho(p^{t+1}, \Phi^t) \le G_\rho(p^{t+1}, \Phi^{t+1}) \le \dots \le C_\rho \tag{1.53}$$

Convém esclarecer que a convergência de iterações para o processo de cálculo de C_ρ segue da desigualdade provada no seguinte teorema (Arimoto (1976)).

Teorema 1.5.2

Seja $p^o \varepsilon P^n$ um dos vetores probabilidade de entrada que maximiza $G_\rho(p)$
Então para algum $\rho \varepsilon (-1, \infty)$,

$$C_\rho - G_\rho(p^{t+1}, \Phi^t) - \frac{\rho}{1+\rho} [C_\rho - G_\rho(p^t, \Phi^t)]$$

$$\le \sum_{i=1}^{n} p_i^o \log \frac{p_i^{t+1}}{p_i^t} \tag{1.54}$$

Garantida a monotonicidade do procedimento iterativo pelo teorema 1.5.2, Arimoto mostrou que:

Teorema 1.5.3

A seqüência $\{G_\rho(p^{t+1}, \Phi^t)\}$ ou $\{G_\rho(p^t, \Phi^t)\}$ definida por (1.51) e (1.52) converge monotonicamente de baixo para C_ρ.

CAPÍTULO 2

LIMITES SUPERIOR E INFERIOR DA CAPACIDADE GENERALIZADA

Arimoto (1972) apresentou os limites superior e inferior da capacidade do canal discreto sem memória, os quais foram melhores que os obtidos anteriormente por Helgert (1967). Neste capítulo são apresentados, pela primeira vez na literatura, os limites superior e inferior da capacidade generalizada, C_ρ. No final deste capítulo apresentamos alguns valores numéricos para três canais diferentes, dos teoremas principais deste trabalho e, também para C_ρ, para cada ρ. Nas conclusões finais, daremos possíveis aplicações dos limites obtidos.

A seguir os cinco teoremas. Os três primeiros dão limites inferiores e os dois últimos dão limites superiores da capacidade generalizada, C_ρ.

Teorema 2.1

$$C_\rho \geq \log n - \frac{1}{\rho} \log \left[\left(\frac{\Delta}{n}\right)^{\frac{1}{1+\rho}} + \left(1 - \frac{\Delta}{n}\right)^{\frac{1}{1+\rho}} \cdot (n-1)^{\frac{\rho}{1+\rho}} \right]^{1+\rho} \qquad (2.1)$$

para $-1 < \rho < \infty$

sendo Δ definido como

$$\Delta = \sum_{j=1}^{m} p(j/i_j) \qquad (2.2)$$

onde, i_j denota um inteiro arbitrariamente escolhido entre 1 e n, correspondentemente a cada j.

Prova

Seja ε um nº arbitrário, tal que $0 \leq \varepsilon \leq 1$ e definimos

$$p_i = \frac{1}{n} \quad , \quad i = 1,2,\ldots,n$$

e

$$\Phi(i/j) = \begin{cases} 1-\varepsilon & \text{se } i = i_j \\ \\ \dfrac{\varepsilon}{n-1} & \text{se } i \neq i_j \end{cases}$$

$$i = 1,2,\ldots,n \qquad\qquad j = 1,2,\ldots,m$$

Agora, por (1.44) e (1.30) temos:

$$C_\rho \geq G_\rho(p,\Phi) = -\frac{1}{\rho}\log\left[\sum_{j=1}^{m}\sum_{i=1}^{n} p_i \, p(j/i)\, \left\{\frac{\Phi(i/j)}{p_i}\right\}^{-\rho}\right]$$

$$-1 < \rho < \infty$$

$$= -\frac{1}{\rho}\log\left[\sum_{j=1}^{m}\sum_{i=1}^{n} \frac{1}{n}\, p(j/i)\, \{\Phi(i/j).n\}^{-\rho}\right]$$

$$= -\frac{1}{\rho}\log\left[\sum_{j=1}^{m}\sum_{i=1}^{n} \frac{1}{n}\, p(j/i)\, \{\Phi(i/j)\}^{-\rho}.n^{-\rho}\right]$$

$$= \log n - \frac{1}{\rho}\log\left[\sum_{j=1}^{m} \frac{p(j/i_j)}{n}\{\Phi(i_j/j)\}^{-\rho} + \right.$$

$$\left. + \sum_{j=1}^{m}\sum_{i\neq i_j}^{n} \frac{p(j/i)}{n}\, \{\Phi(i/j)\}^{-\rho}\right]$$

$$= \log n - \frac{1}{\rho} \log [\frac{\Delta}{n} (1 - \varepsilon)^{-\rho} + (1 - \frac{\Delta}{n}) (\frac{\varepsilon}{n-1})^{-\rho}]$$

Então

$$C_\rho \geq \log n - \frac{1}{\rho} \log [(\frac{\Delta}{n}) . (1 - \varepsilon)^{-\rho} + (1 - \frac{\Delta}{n}) (\frac{\varepsilon}{n-1})^{-\rho}] \tag{2.3}$$

Maximizando o lado direito de (2.3) em relação a $0 \leq \varepsilon \leq 1$, obtemos

$$\varepsilon = \frac{1}{[\frac{\Delta}{(n-\Delta)(n-1)^\rho}]^{1/1+\rho} + 1}$$

De fato, seja

$$F(\varepsilon) = \log n - \frac{1}{\rho} \log [(\frac{\Delta}{n}) (1 - \varepsilon)^{-\rho} + (1 - \frac{\Delta}{n}) (\frac{\varepsilon}{n-1})^{-\rho}]$$

Como

$$\frac{\partial F(\varepsilon)}{\partial \varepsilon} = - \frac{1}{\rho} [\frac{\frac{1}{n} \{\frac{\Delta\rho(1-\varepsilon)^{\rho-1}}{(1-\varepsilon)^{2\rho}} - \frac{(n-\Delta)(n-1)^\rho . \rho\varepsilon^{\rho-1}}{\varepsilon^{2\rho}}\}}{\frac{1}{n} \{\frac{\Delta\varepsilon^\rho + (n-\Delta)(n-1)^\rho . (1-\varepsilon)^\rho}{\varepsilon^\rho (1-\varepsilon)^\rho}\}}] = 0$$

Simplificando, chegamos a

$$\Delta\varepsilon^{\rho+1} - (n-\Delta)(n-1)^\rho (1-\varepsilon)^{\rho+1} = 0$$

Portanto $\varepsilon = \dfrac{1}{\left[\dfrac{\Delta}{(n-\Delta)(n-1)^{\rho}}\right]^{1/1+\rho} + 1}$

ou

$$\varepsilon = \frac{[(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}}}{\Delta^{\frac{1}{1+\rho}} + [(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}}}$$

Substituindo ε em (2.3) temos:

$$C_{\rho} \geq \log n - \frac{1}{\rho}\log\left[\frac{\Delta}{n\left(1 - \dfrac{[(n-\Delta)(n-1)^{\rho}]^{1/1+\rho}}{\Delta^{\frac{1}{1+\rho}} + [(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}}}\right)^{\rho}}\right.$$

$$\left. + \left(1 - \frac{\Delta}{n}\right)(n-1)^{\rho}\left(\frac{[(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}} + \Delta^{\frac{1}{1+\rho}}}{[(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}}}\right)^{\rho}\right]$$

$$= \log n - \frac{1}{\rho}\log\left[\frac{\Delta\left(\Delta^{\frac{1}{1+\rho}} + [(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}}\right)^{\rho}}{n\left(\Delta^{\frac{\rho}{1+\rho}}\right)} + \right.$$

$$\left. \frac{(n-\Delta)(n-1)^{\rho}\left([(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}} + \Delta^{\frac{1}{1+\rho}}\right)^{\rho}}{n[(n-\Delta)(n-1)^{\rho}]^{\frac{\rho}{1+\rho}}}\right]$$

$$= \log n - \frac{1}{\rho}\log\left[\frac{1}{n}\left\{\Delta^{\frac{1}{1+\rho}}\left(\Delta^{\frac{1}{1+\rho}} + [(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}}\right)^{\rho} + \right.\right.$$

$$\left.\left. (n-\Delta)^{\frac{1}{1+\rho}}(n-1)^{\frac{\rho}{1+\rho}}\left(\Delta^{\frac{1}{1+\rho}} + [(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}}\right)^{\rho}\right\}\right]$$

$$= \log n - \frac{1}{\rho} \log\ [(\tfrac{1}{n})(\Delta^{\frac{1}{1+\rho}} + [(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}})^{\rho}(\Delta^{\frac{1}{1+\rho}} + [(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}})$$

$$= \log n - \frac{1}{\rho} \log\ [(\tfrac{1}{n}).(\Delta^{\frac{1}{1+\rho}} + [(n-\Delta)(n-1)^{\rho}]^{\frac{1}{1+\rho}})^{1+\rho}]$$

Portanto

$$C_\rho \geq \log n - \frac{1}{\rho} \log\ [(\tfrac{\Delta}{n})^{\frac{1}{1+\rho}} + (1-\tfrac{\Delta}{n})^{\frac{1}{1+\rho}}\ (n-1)^{\frac{\rho}{1+\rho}}]^{1+\rho}$$

<u>Teorema 2.2</u>

$$C_\rho \geq \log\ [\sum_{i=1}^{n}\ \{\sum_{j=1}^{m}\ p(j/i)^{\frac{1}{1+\rho}}\ (\sum_{i=1}^{n}\ p(j/i)^{\frac{1}{1+\rho}})^{\rho}\}^{-\frac{1}{\rho}}] \qquad (2.4)$$

$$-1 < \rho < \infty$$

<u>Prova</u>

Por (1.36) e (1.30), temos

$$G_\rho \geq G_\rho(\Phi) = \max_{p\ \varepsilon\ P^n} G_\rho(p,\ \Phi)$$

isto é

$$C_\rho \geq \log\ [\sum_{i=1}^{n}\ \{\sum_{j=1}^{m}\ p(j/i)\ [\Phi(i/j)]^{-\rho}\}^{-\frac{1}{\rho}}] \qquad (2.5)$$

Fazendo $p_i = \frac{1}{n}$ $\qquad i = 1,2,\ldots,n$ $\qquad$ em

$$\Phi(i/j) = \frac{p_i\ p(j/i)^{\frac{1}{1+\rho}}}{\sum_{k=1}^{n} p_k\ p(j/k)^{\frac{1}{1+\rho}}}$$

temos

$$\Phi(i/j) = \frac{p(j/i)^{\frac{1}{1+\rho}}}{\sum_{k=1}^{n} p(j/k)^{\frac{1}{1+\rho}}} \tag{2.6}$$

Substituindo este valor de $\Phi(i/j)$ de (2.6) em (2.5) temos:

$$C_\rho \geq \log \left[\sum_{i=1}^{n} \left\{ \sum_{j=1}^{m} p(j/i) \left(\frac{p(j/i)^{\frac{1}{1+\rho}}}{\sum_{k=1}^{n} p(j/k)^{\frac{1}{1+\rho}}} \right)^{-\rho} \right\}^{-\frac{1}{\rho}} \right]$$

Então

$$C_\rho \geq \log \left[\sum_{i=1}^{n} \left\{ \sum_{j=1}^{m} p(j/i)^{\frac{1}{1+\rho}} \left[\sum_{i=1}^{n} p(j/i)^{\frac{1}{1+\rho}} \right]^{\rho} \right\}^{-\frac{1}{\rho}} \right] .$$

o que prova o teorema.

<u>Teorema 2.3</u>

$$C_\rho \geq \frac{1+\rho}{\rho} \log n - \frac{1}{\rho} \log \left[\sum_{j=1}^{m} \left\{ \sum_{i=1}^{n} p(j/i)^{\frac{1}{1+\rho}} \right\}^{1+\rho} \right] \tag{2.7}$$

$$-1 < \rho < \infty$$

<u>Prova</u> Por (1.30) temos:

$$C_\rho \geq -\frac{1}{\rho} \log \left[\sum_{j=1}^{m} \sum_{i=1}^{n} p_i \, p(j/i) \left\{ \frac{\Phi(i/j)}{p_i} \right\}^{-\rho} \right]$$

$$-1 < \rho < \infty$$

Seja $p_i = \frac{1}{n}$ $\quad i = 1,2,\ldots,n$

então, como $\Phi(i/j) = \dfrac{p(j/i)^{\frac{1}{1+\rho}}}{\sum_{k=1}^{n} p(j/k)^{\frac{1}{1+\rho}}}$ temos

$$C_\rho \geq -\frac{1}{\rho}\log\left[\sum_{j=1}^{m}\frac{1}{n}\sum_{i=1}^{n} p(j/i)\left\{\frac{p(j/i)^{\frac{1}{1+\rho}}}{\sum_{k=1}^{n} p(j/k)^{\frac{1}{1+\rho}}}\cdot n\right\}^{-\rho}\right]$$

$$C_\rho \geq -\frac{1}{\rho}\log\left[\sum_{j=1}^{m}\sum_{i=1}^{n}\frac{1}{n^{1+\rho}}\, p(j/i)\,\{p(j/i)^{\frac{1}{1+\rho}}\}^{-\rho}\cdot\left(\sum_{k=1}^{n} p(j/k)^{\frac{1}{1+\rho}}\right)^{\rho}\right]$$

$$= -\frac{1}{\rho}\log\left[\frac{1}{n^{1+\rho}}\sum_{j=1}^{m}\sum_{i=1}^{n} p(j/i)^{\frac{1}{1+\rho}}\left\{\sum_{k=1}^{n} p(j/k)^{\frac{1}{1+\rho}}\right\}^{\rho}\right]$$

$$= \frac{1+\rho}{\rho}\log n - \frac{1}{\rho}\log\left[\sum_{j=1}^{m}\left\{\sum_{i=1}^{n} p(j/i)^{\frac{1}{1+\rho}}\right\}^{1+\rho}\right]$$

o que prova o teorema.

Teorema 2.4

$$C_\rho \leq \begin{cases} \log m - \dfrac{1+\rho}{\rho}\log\left\{\max\limits_{i}\left(\sum_{j=1}^{m} p(j/i)^{\frac{1}{1+\rho}}\right)\right\}; & -1 < \rho < 0 \\[2ex] \log m - \dfrac{1+\rho}{\rho}\log\left\{\min\limits_{i}\left(\sum_{j=1}^{m} p(j/i)^{\frac{1}{1+\rho}}\right)\right\}; & \rho > 0 \end{cases}$$

Prova

Seja $r_j^o = \sum_{i=1}^{n} p_i^o\, p(j/i)^{\frac{1}{1+\rho}}$

Considere

$$\sum_{j=1}^{m} (r_j^o)^{1+\rho} = m \sum_{j=1}^{m} \left[\frac{1}{m}(r_j^o)^{1+\rho}\right]$$

$$= m \sum_{j=1}^{m} [\frac{1}{m} (\sum_{i=1}^{n} p_i^o \; p(j/i)^{\frac{1}{1+\rho}})^{1+\rho}]$$

$$\leq m \, [\frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{n} p_i^o \; p(j/i)^{\frac{1}{1+\rho}}]^{1+\rho} \quad \text{para} \quad 0 < 1+\rho < 1$$

Gallager (1968))

onde $p^o = (p_1^o, p_2^o, \ldots, p_n^o) \in \overline{P}^n$ é a distribuição de probabilidade que atinge a capacidade de ordem ρ, isto é

$$\log [m \sum_{j=1}^{m} \frac{1}{m} (r_j^o)^{1+\rho}]$$

$$\leq \log [m \, (\frac{1}{m} \sum_{j=1}^{m} \sum_{i=1}^{n} p_i^o \; p(j/i)^{\frac{1}{1+\rho}})^{1+\rho}] \;, \quad -1 < \rho < 0$$

$$= \log [m . (\frac{1}{m})^{1+\rho} (\sum_{j=1}^{m} \sum_{i=1}^{n} p_i^o \; p(j/i)^{\frac{1}{1+\rho}})^{1+\rho}]$$

$$= -\rho \log m + (1+\rho) \log [\sum_{j=1}^{m} \sum_{i=1}^{n} p_i^o \; p(j/i)^{\frac{1}{1+\rho}}]$$

$$\leq -\rho \log m + (1+\rho) \log [\max_i (\sum_{j=1}^{m} p(j/i)^{\frac{1}{1+\rho}})], \; -1 < \rho < 0.$$

Pois

$$\sum_{i=1}^{n} p_i^o \sum_{j=1}^{m} p(j/i)^{\frac{1}{1+\rho}} \leq \max_i (\sum_{j=1}^{m} p(j/i)^{\frac{1}{1+\rho}}$$

agora, para $\rho < 0$, $-\frac{1}{\rho} > 0$ e temos

$$-\frac{1}{\rho}\{m\log[\frac{1}{m}(\sum_{j=1}^{m} r_j^o)^{1+\rho}]\} \leq \log m - \frac{1+\rho}{\rho}\log\{\max_i(\sum_{j=1}^{m} P(j/i)^{\frac{1}{1+\rho}})$$

Portanto

$$C_\rho \leq \log m - \frac{1+\rho}{\rho}\log\{\max_i(\sum_{j=1}^{m} P(j/i)^{\frac{1}{1+\rho}})\}$$

$$-1 < \rho < 0 \qquad\qquad (\text{por } (2.2))$$

Similarmente, usando o fato que

$$\sum_{i=j}^{n} p_i^o \sum_{j=1}^{m} p(j/i)^{\frac{1}{1+\rho}} \geq \min_i(\sum_{j=1}^{m} P(j/i)^{\frac{1}{1+\rho}})$$

Obtemos

$$C_\rho \leq \log m - \frac{1+\rho}{\rho}\log\{\min_i(\sum_{j=1}^{m} P(j/i)^{\frac{1}{1+\rho}})\}, \quad \rho > 0$$

<u>Teorema 2.5</u>

$$C_\rho \leq \log n - \frac{1}{\rho}\log\{\max_i \Big| \sum_{j=1}^{m} P(j/i)^{\frac{1}{1+\rho}}(\sum_{k=1}^{n} P(j/k)^{\frac{1}{1+\rho}})^{\rho}]\}$$

$$-1 < \rho \leq 0$$

<u>Prova</u>

Segue imediato do Lema 1.4.2 tomando

$$p_i = \frac{1}{n} \qquad i = 1,2,\ldots,n$$

CONCLUSÕES

Arimoto (1975) escreveu o teorema fundamental da teoria da informação (teorema de codificação aleatória), provado por Gallager (1965) e seu contrário por Arimoto (1973), em função da capacidade generalizada, C_ρ (Teorema 1.3.4). Em outras palavras, a conhecida função exponencial de codificação aleatória $E_0(\rho,p)$ foi escrita por Arimoto (1975) em termos da informação mútua generalizada de ordem ρ (em função da entropia de ordem α, $\alpha = \frac{1}{1+\rho}$).

Quanto a este trabalho, servem os dois motivos:

1. Generalizar os limites superior e inferior, obtidos por Arimoto (1972) para os canais discretos sem memória.
2. Uma vez obtidos os limites, por conhecimento somente da matriz do canal, podemos saber através do teorema 1.3.4, os limites superior e inferior da probabilidade de erro.

Estes limites obtidos podem ser usados posteriormente para limitar outras funções importantes da teoria da informação, tais como: "Reliability - Rate Functions", "Exponent Functions for Constrained Channels", "Source Code Reliability Rate Functions", "Omura's Exponent Functions etc., (Arimoto (1976))

Pelos três exemplos particulares de matrizes de canais consideradas no anexo, verificamos que em determinados exemplos alguns teoremas fornecem melhor aproximação do que outros.

Fica, em aberto, como estudo posterior, verificar e provar teoricamente, as possíveis comparações entre os limites obtidos. E também, em aberto, os resultados correspondentes para $\rho>0$ no teorema 2.5.

EXEMPLOS NUMÉRICOS

$$\text{Matriz Canal} - \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.7 & 0.1 & 0.2 \\ 0.5 & 0.05 & 0.45 \end{pmatrix}$$

ρ	C_ρ	Teorema 2.2	Teorema 2.3	Teorema 2.4	Teorema 2.5	Distribuição Maximiza
0.9	0.4391	0.4383	0.4368	1.0132	0.6933	(0.253, 0.431, 0.316)
0.8	0.3663	0.3657	0.3653	0.9424	0.4433	(0.366, 0.271, 0.363)
0.7	0.3186	0.3101	0.3074	0.8602	0.4378	(0.470, 0.107, 0.423)
0.6	0.2857	0.2645	0.2605	0.7755	0.3915	(0.526, 0.002, 0.472)
0.5	0.2575	0.2272	0.2234	0.6960	0.3383	(0.516, 0.000, 0.484)
0.4	0.2325	0.1974	0.1940	0.6253	0.2904	(0.510, 0.000, 0.490)
0.3	0.2107	0.1734	0.1706	0.5640	0.2506	(0.506, 0.000, 0.494)
0.2	0.1918	0.1540	0.1518	0.5115	0.2183	(0.504, 0.000, 0.496)
0.1	0.1756	0.1383	0.1365	0.4666	0.1921	(0.503, 0.000, 0.497)
0.0	0.1616	0.1252	0.1236	0.4287	0.1708	(0.502, 0.000, 0.498)
0.1	0.1495	0.1143	0.1132	0.3950		(0.501, 0.000, 0.499)
0.2	0.1389	0.1051	0.1041	0.3662		(0.501, 0.000, 0.499)
0.3	0.1296	0.0972	0.0944	0.3411		(0.500, 0.000, 0.500)
0.4	0.1214	0.0904	0.0897	0.3190		(0.500, 0.000, 0.500)
0.5	0.1141	0.0844	0.0839	0.2995		(0.500, 0.000, 0.500)
0.6	0.1076	0.0791	0.0787	0.2822		(0.500, 0.000, 0.500)
0.7	0.1018	0.0745	0.0741	0.2666		(0.500, 0.000, 0.500)
0.8	0.0965	0.0704	0.0701	0.2527		(0.500, 0.000, 0.500)
0.9	0.0917	0.0667	0.0664	0.2400		(0.500, 0.000, 0.500)
1.0	0.0874	0.0634	0.0631	0.2286		(0.500, 0.000, 0.500)
2.0	0.0592	0.0422	0.0421	0.1638		(0.499, 0.002, 0.499)
3.0	0.0466	0.0316	0.0315	0.1280		(0.498, 0.004, 0.498)
4.0	0.0356	0.0252	0.0252	0.1048		(0.496, 0.008, 0.496)
5.0	0.0296	0.0210	0.0210	0.0887		(0.493, 0.014, 0.493)
6.0	0.0252	0.0180	0.0180	0.0769		(0.489, 0.022, 0.489)
7.0	0.0219	0.0157	0.0157	0.0678		(0.484, 0.032, 0.484)
8.0	0,0192	0,0140	0.0140	0.0607		(0.478, 0.044, 0.478)
9.0	0.0171	0.0126	0.0126	0.0549		(0.469, 0.062, 0.469)
10.0	0.0152	0.0114	0.0114	0.0501		(0.459, 0.082, 0.459)

<u>Observação 1</u> - Arimoto (1972) e Helgert (1967) também obtiveram os valores numéricos usando este canal para o caso particular $\rho = 0$.

<u>Observação 2</u> - O limite superior obtido por Helgert (1967) é 0,55 bits para $\rho = 0$, mas em nosso caso temos 0,4287 bits no teorema 2.4. Enquanto no teorema 2.5 é ainda melhor, e é exatamente o obtido por Arimoto (1972); 0,1708 bits.

$$\text{Matriz Canal} - \begin{pmatrix} 1/2 & 1/3 & 1/6 \\ 1/3 & 1/2 & 1/6 \\ 1/3 & 1/6 & 1/2 \end{pmatrix}$$

ρ	C_ρ	Teorema 2.2	Teorema 2.3	Teorema 2.4	Teorema 2.5	Distribuição Maximiza
0.9	0.4766	0.4766	0.4766	0.4766	0.5006	(0.328, 0.333, 0.339)
0.8	0.3809	0.3805	0.3797	0.3809	0.5173	(0.284, 0.340, 0.376)
0.7	0.3112	0.3088	0.3064	0.3112	0.4982	(0.230, 0.357, 0.413)
0.6	0.2603	0.2547	0.2517	0.2603	0.4423	(0.180, 0.381, 0.439)
0.5	0.2224	0.2140	0.2111	0.2224	0.3795	(0.138, 0.405, 0.457)
0.4	0.1934	0.1831	0.1807	0.1934	0.3230	(0.105, 0.425, 0.470)
0.3	0.1708	0.1592	0.1573	0.1708	0.2762	(0.079, 0.443, 0.478)
0.2	0.1527	0.1405	0.1390	0.1527	0.2383	(0.059, 0.456, 0.485)
0.1	0.1380	0.1255	0.1243	0.1380	0.2079	(0.045, 0.466, 0.489)
0.0	0.1257	0.1133	0.1123	0.1257	0.1832	(0.036, 0.473, 0.491)
0.1	0.1154	0.1032	0.1025	0.1155		(0.030, 0.477, 0.493)
0.2	0.1065	0.0947	0.0941	0.1068		(0.027, 0.479, 0.494)
0.3	0.0989	0.0875	0.0870	0.0992		(0.025, 0.481, 0.494)
0.4	0.0923	0.0813	0.0809	0.0927		(0.024, 0.481, 0.495)
0.5	0.0865	0.0759	0.0755	0.0870		(0.024, 0.481, 0.495)
0.6	0.0814	0.0711	0.0708	0.0819		(0.025, 0.480, 0.495)
0.7	0.0768	0.0669	0.0667	0.0773		(0.026, 0.479, 0.495)
0.8	0.0727	0.0632	0.0630	0.0733		(0.027, 0.478, 0.495)
0.9	0.0690	0.0600	0.0597	0.0696		(0.028, 0.477, 0.495)
1.0	0.0656	0.0569	0.0567	0.0663		(0.030, 0.476, 0.494)
2.0	0.0440	0.0379	0.0378	0.0449		(0.054, 0.455, 0.491)
3.0	0.0328	0.0283	0.0283	0.0339		(0.087, 0.426, 0.487)
4.0	0.0260	0.0227	0.0226	0.0272		(0.129, 0.389, 0.482)
5.0	0.0213	0.0189	0.0189	0.0228		(0.177, 0.350, 0.473)
6.0	0.0180	0.0162	0.0162	0.0195		(0.213, 0.330, 0.457)
7.0	0.0156	0.0141	0.0141	0.0171		(0.235, 0.323, 0.442)
8.0	0.0137	0.0126	0.0126	0.0152		(0.253, 0.321, 0.426)
9.0	0.0122	0.0113	0.0113	0.0137		(0.268, 0.321, 0.411)
10.0	0.0109	0.0103	0.0103	0.0125		(0.283, 0.323, 0.394)

$$\text{Matriz Canal} = \begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.7 & 0.0 & 0.1 & 0.2 \\ 0.0 & 0.5 & 0.3 & 0.2 \end{pmatrix}$$

ρ	C_ρ	Teorema 2.2	Teorema 2.3	Teorema 2.4	Teorema 2.5	Distribuição Maximiza
-0.9	0.8718	0.8717	0.8709	1.4282	1.0141	(0.277, 0.368, 0.355)
-0.8	0.8119	0.8111	0.8087	1.3575	0.9739	(0.252, 0.376, 0.372)
-0.7	0.7600	0.7557	0.7497	1.2752	0.9614	(0.200, 0.396, 0.404)
-0.6	0.7218	0.7084	0.6976	1.1905	0.9509	(0.128, 0.422, 0.450)
-0.5	0.6972	0.6676	0.6523	1.1110	0,9323	(0.045, 0.450, 0.505)
-0.4	0.6833	0.6317	0.6130	1.0403	0.9056	(0.000, 0.466, 0.534)
-0.3	0.6713	0.5998	0.5789	0.9791	0.8736	(0.000, 0.470, 0.530)
-0.2	0.6599	0.5713	0.5~~7~~493	0.9266	0.8390	(0.000, 0.474, 0.526)
-0.1	0.6490	0.5456	0.5233	0.8817	0.8039	(0.000, 0.477, 0.523)
0.0	0.6385	0.5226	0.5013	0.8430	0.7679	(0.000, 0.480, 0.520)
0.1	0.6282	0.5017	0.4804	0.8100		(0.000, 0.483, 0.517)
0.2	0.6182	0.4828	0.4623	0.7813		(0.000, 0.485, 0.515)
0.3	0.6084	0.4656	0.4462	0.7561		(0.000, 0.488, 0.512)
0.4	0.5986	0.4499	0.4315	0.7341		(0.000, 0.490, 0.510)
0.5	0.5891	0.4354	0.4182	0.7146		(0.000, 0.492, 0.508)
0.6	0.5796	0.4222	0.4060	0.6972		(0.000, 0.494, 0.506)
0.7	0.5702	0.4099	0.3947	0.6817		(0.000, 0.495, 0.505)
0.8	0.5608	0.3985	0.3843	0.6677		(0.000, 0.497, 0.503)
0.9	0.5516	0.3878	0.3746	0.6551		(0.000, 0.499, 0.501)
1.0	0.5424	0.3778	0.3655	0.6437		(0.000, 0.500, 0.500)
2.0	0.4555	0.3034	0.2972	0.5694		(0.000, 0.513, 0.487)
3.0	0.3801	0.2543	0.2511	0.5313		(0.000, 0.525, 0.475)
4.0	0.3187	0.2180	0.2162	0.5083		(0.000, 0.541, 0.459)
5.0	0.2706	0.1895	0.1885	0.4928		(0.000, 0.561, 0.439)
6.0	0.2335	0.1666	0.1661	0.4817		(0.000, 0.585, 0.415)
7.0	0.2048	0.1479	0.1476	0.4734		(0.000, 0.611, 0.389)
8.0	0.1824	0.1325	0.1323	0.4670		(0.000, 0.637, 0.363)
9.0	0.1644	0.1196	0.1195	0.4618		(0.001, 0.659, 0.340)
10.0	0.1497	0.1088	0.1087	0.4575		(0.001, 0.678, 0.321)

BIBLIOGRAFIA

ARIMOTO, S. (1972). An algorithm for computing the capacity of arbitrary discrete memoryless channels. IEEE Transactions on Information Theory. Vol. IT-18, pp. 14-20.

ARIMOTO, S. (1973). On the converse to the coding theorem for discrete memoryless channels. IEEE Trans. on Information Theory. Vol. IT-19, pp. 357-359.

ARIMOTO, S. (1975). "Information measures and capacity of order α for discrete memoryless channels" presented at the Colloquium on Information Theory. Kesthely Hungary, pp. 41-52.

ARIMOTO, S. (1976). Computation of randon coding exponent functions. IEEE Trans. on Informations Theory. Vol. IT-22, pp. 665-671.

ASH, R.B. (1965). Information Theory. Interscience Publishers, New York.

BLAHUT, R.E. (1972). Computation of channel capacity and rate - distortion functions. IEEE Trans. on Information Theory. Vol. IT-18, pp. 460-473.

GALLAGER, R.G. (1965). A simple derivation of coding theorem and some applications. IEEE Trans. on Information Theory. Vol. IT-11, pp. 3-18.

GALLAGER, R.G. (1968). Information Theory and Reliable Comunication. New York, Wiley.

GUERRA, F. (1980). Entropias Generalizadas e Capacidade do Canal Discreto Sem Memória. Dissertação de Mestrado - UFSC.

HELGERT, H.J. (1967). On a Bound for Channel Capacity. IEEE Trans. on Inform. Theory. IT-13, pp. 124-126.

JELINEK, F. (1968). Probabilistic Information Theory. New York, McGraw-Hill.

OMURA, J.K. (1973). A coding theorem for discrete-time sources. IEEE Trans. on Information Theory. Vol. IT-19, pp. 490-498.

RENYI, A. (1961). On measures of entropy and information. Vol. I, pp. 547-561 in "Proc. of the 4th Berkeley Symposium Mathematical Statistics and Probability". Univ. of California Press, Berkeley.

SHANNON, C.E. (1948). A mathematical theory of communication. Bell System Technical Journal. Vol. 27, pp. 379-423 and 623-656.